Simulated Anneal Method for the Determination of Spin Hamiltonian Parameters from ESR Data

Heynderickx, I.; Raedt, H. De; Schoemaker, D.

Published in:
Journal of Magnetic Resonance %281969%29

DOI:
10.1016/0022-2364(86)90367-7

Publication date:
1986

JOURNAL OF MAGNETIC RESONANCE 70, 134–139 (1986)

Simulated Anneal Method for the Determination of Spin Hamiltonian Parameters from ESR Data

I. HEYNDERICKX, H. DE RAEDT, AND D. SCHOEMAKER

Physics Department, University of Antwerp (U.I.A.), B-2610 Wilrijk-Antwerpen, Belgium

Received April 1, 1986; revised June 6, 1986

The simulated anneal procedure based on the Metropolis Monte Carlo algorithm is presented as a new numerical method for the analysis of electron-spin-resonance spectra by the spin Hamiltonian formalism. It gives an alternative way for finding the global minimum of a function, which describes the difference for each of the resonance fields between the theoretically calculated energy difference and the experimentally measured microwave energy. Compared to the iterative procedures commonly used, it has the advantage of not getting stuck in local minima of the parameter space. This implies that the procedure can be started with arbitrary initial parameters. The usefulness of the method is illustrated by its application to the analysis of four defects in alkali halides doped with heavy metal ions.

INTRODUCTION

Electron-spin-resonance spectra are usually fitted to an effective spin Hamiltonian (*1*). In the case of an effective spin $S = \frac{1}{2}$ and nuclear spin $I = \frac{1}{2}$, this Hamiltonian has the form

$$\frac{\mathscr{H}}{g_0\mu_B} = \mathbf{H}\cdot\mathbf{g}\cdot\mathbf{S}+\mathbf{S}\cdot\mathbf{A}\cdot\mathbf{I}, \qquad [1]$$

with g_0 the free electron g value and μ_B the Bohr magneton. The tensor $\mathbf{g}$ describes the Zeeman interaction of $\mathbf{S}$ with the external static magnetic field $\mathbf{H}$. The hyperfine tensor $\mathbf{A}$ determines the interaction of $\mathbf{S}$ with $\mathbf{I}$. The principal values of the $\mathbf{g}$ and $\mathbf{A}$ tensor are obtained by fitting the theoretical spectrum, predicted by [1], to the experimental spectrum. The basic information provided by the latter consists of a set of resonance fields $\{\mathbf{H}_l;\ l = 1, \ldots, L\}$ measured at the corresponding microwave frequencies $\{\nu_l;\ l = 1, \ldots, L\}$. As is well known, this information is independent of the sign of the hyperfine parameters such that only their absolute value can be determined. The theoretical spectrum can be fitted to this experimental information by minimizing the function

$$f(\mathbf{g},\mathbf{A}) \equiv \sum_{l=1}^{L} |E_{n_l}(\mathbf{g},\mathbf{A},\mathbf{H}_l) - E_{n'_l}(\mathbf{g},\mathbf{A},H_l) - h\nu_l|, \qquad [2]$$

where $E_i(\mathbf{g}, \mathbf{A}, \mathbf{H}_l)$ denotes the ith eigenvalue of Hamiltonian [1] for a given choice of $\mathbf{g}$, $\mathbf{A}$ and $\mathbf{H}_l$, and h is Planck's constant. For each term in the sum over the exper-

0022-2364/86 $3.00

imentally measured magnetic fields, the indices n_I and n_I' are determined by the following rules:

(1) the corresponding energy difference has to approach the microwave energy as well as possible in order to fulfill the resonance condition;

(2) the transition between these energy levels has to be allowed according to the selection rules $\Delta m_S = 1$ and $\Delta m_I = 0$ (*1*). This is checked by calculating the intensity of the transition from the sum of the $\Delta m_S = 1$, $\Delta m_I = 0$ contributions for the eigenstates which are mixtures of the $|m_S, m_I>$ states.

Obviously the choice of function [2] is not unique for the determination of the spin Hamiltonian parameters. Taking, for example, the sum over the nth power of the absolute values in function [2] is a trivial modification resulting in a minimum for the same set of parameters. In fact these parameters can also be determined from the minimization of the difference between experimentally measured and theoretically calculated quantities such as the resonance magnetic fields or the intensity of the transition for different magnetic fields.

For spin Hamiltonian [1] with a hyperfine interaction small compared to the Zeeman interaction, perturbative expressions for the spin Hamiltonian parameters (*1*) are used for each direction of the static magnetic field **H** with respect to the center axes in order to determine the minimum of $f(\mathbf{g}, \mathbf{A})$. Exact expressions for the g and A parameters are given by the Breit-Rabi formulae (*2*) in the case of isotropic defects possessing a small or a large hyperfine interaction. For the analysis of an ESR spectrum in a general orientation of the static magnetic field with respect to the center axes of a system with axial or orthorhombic symmetry and characterized by a large hyperfine interaction, numerical diagonalization of the Hamiltonian matrix is necessary. Consequently minimizing functions such as $f(\mathbf{g}, \mathbf{A})$ require the solution of a highly nonlinear problem. To do so one normally uses an iterative method (*3–6*) such as the least-squares fit (*5*) or the gradient projection (*6*) method. For all these methods convergence is never guaranteed since the search can always get stuck in a local minimum. In order to find the global minimum by an iterative method one has to find initial values which correspond approximately to the minimum of this function. Such values are usually derived from a second order perturbation solution of [1] (*4*).

In the present paper we report the use of the simulated anneal (SA) technique based on the Metropolis Monte Carlo method for the minimization of $f(\mathbf{g}, \mathbf{A})$. The usefulness of this method for finding the minimum of a given nonlinear function depending on many parameters in a complicated manner has already been demonstrated by Kirkpatrick *et al.* (*7*) in the optimization of computer design. Characteristic for this method is that steps in the multidimensional parameter space are such that transitions out of a local minimum are possible. Recently a different simulated anneal method has been proposed (*8*). In this approach the thermal annealing is accomplished by solving the equations of motion by means of the molecular dynamics technique.

THE SIMULATED ANNEAL METHOD

In classical statistical mechanics the mean value of a function f for a system in thermal equilibrium at a given temperature T is given by the relation

$$\langle f \rangle = \frac{\int f(S) e^{-\beta E(S)} dS}{\int e^{-\beta E(S)} dS}, \qquad [3]$$

in which $E(S)$ is the energy of the system in its state S and $\beta = 1/k_B T$ with k_B Boltzmann's constant. From [3] it follows immediately that as $T \to 0$ ($\beta \to \infty$) the dominant contribution is given by states S with minimal energy. Assuming there is a unique ground state one has

$$\lim_{\beta \to \infty} \langle f \rangle = f(S_0), \qquad [4]$$

where S_0 is the state that possesses the minimal energy E_0. The SA method is based on the observation that by cooling the system, that is, by increasing β, the system will approach the state S_0. Consequently it can be used for the minimization of $E(S)$. If we are interested in a particular classical statistical model $E(S)$ must be the energy of the state S. For our problem we need only replace $E(S)$ by $f(\mathbf{g}, \mathbf{A})$ and regard $\mathbf{g}$ and $\mathbf{A}$ as the degrees of freedom in this artificial system.

So far we have assumed that we can actually calculate quantities such as [3], but this is by no means a trivial problem. Fortunately we can employ the Metropolis Monte Carlo method (*9, 10*), which generates states with a probability density of $\pi(S)$ where

$$\pi(S) = \frac{e^{-\beta E(S)}}{\int e^{-\beta E(S)} dS}, \qquad [5]$$

so that

$$\langle f \rangle = \frac{1}{\#S} \sum_{\{S\}} f(S) \pi(S). \qquad [6]$$

In this algorithm a trial state S' is generated by application of a matrix M, consisting of random numbers, to the old state S and $\pi(S')$ is calculated. If $\pi(S')/\pi(S) > 1$ the state S' is accepted as the new state. On the other hand if $\pi(S')/\pi(S) \leqslant 1$ the state S' is accepted when $\pi(S')/\pi(S) \geqslant r$, where $0 < r < 1$ is a random number drawn from a uniform distribution. Repeating this procedure over and over again it follows that the states generated by this algorithm will be distributed according to $\pi(S)$ on condition that the matrix $M = (m_{ij})$ is positive and symmetric, which means

$$m_{ij} > 0 \qquad \forall i, j, \qquad [7a]$$

and

$$m_{ij} = m_{ji} \qquad \forall i, j. \qquad [7b]$$

In practice this algorithm is implemented as follows. A trial set of g and A parameters is generated by two random numbers, one for selecting the parameter to be changed and one for determining the trial value for this parameter. The value of function [2] for the trial set is evaluated and eventually accepted as a new set of parameters following the prescription of the Metropolis procedure. The condition [7a] is fulfilled by introducing a maximal step size equal to half of the allowed interval for each of the param-

eters, so that the parameter selected by the random number can obtain any value within this interval by a single step. Moreover in order to make the matrix M symmetric, periodic boundary conditions are involved for steps near the edges of the parameter intervals.

As explained above, the minimal value of $f(\mathbf{g}, \mathbf{A})$ is obtained by increasing β. In our calculations we have chosen to increase β after a number of steps equal to the number of parameters to be changed. The percentage of accepted Monte Carlo steps is controlled by the initial and final value of β and its increment. Indeed, for larger values of β the distribution around the minimum of $f(\mathbf{g}, \mathbf{A})$ is smaller and steps that result in an increase of $f(\mathbf{g}, \mathbf{A})$ are less probable.

Finally we have to remark that by calculating the value of function [2] it is necessary to introduce a supplementary physical constraint because it is not permitted to select the same pair of levels (n_l, n_l') for magnetic fields belonging to the same orientation. Otherwise the function could be minimized by taking energy levels which are a constant function of the magnetic field, in contradiction to the behavior of the Zeeman interaction.

APPLICATION

The computer program as described in the previous section was applied to four defects which have already been studied at X-band frequencies ($\nu \simeq 9.3$ GHz) in alkali halides doped with heavy metal ions, namely the $Sn^+(1)$ defect in KCl (*11*), the Tl^0(ortho, I) defect in NaCl (*12*), the $Tl_2^+\langle 110\rangle$ defect in KCl (*13*) and the Tl^+-perturbed $Tl^0(1)$ defect in KCl (*14*). The $Sn^+(1)$ defect is a ($S = \frac{1}{2}$, $I = \frac{1}{2}$) system described by a small hyperfine interaction comparable to the Zeeman interaction, while for the Tl^0(ortho, I) defect the ($S = \frac{1}{2}$, $I = \frac{1}{2}$) system is characterized by a large hyperfine interaction. For the $Tl_2^+\langle 110\rangle$ and the Tl^+-perturbed $Tl^0(1)$ defect the $S = \frac{1}{2}$ electron spin interacts with two equivalent $I = \frac{1}{2}$ nuclei. For the $Tl_2^+\langle 110\rangle$ defect the principal axes of the **g** and **A** tensors coincide, while for the Tl^+-perturbed $Tl^0(1)$ defect the z axes of the **A** tensor are tilted away from the z axis of the **g** tensor over angles $+\delta$ and $-\delta$ in the (xz) plane, defined by the principal axes of the **g** tensor.

In Table 1 we present the results of both the iterative and the SA calculations. From the values of the parameters determined iteratively, it follows that the hyperfine parameters of these four defects are spread over a big range. Nevertheless, each of the defects is analyzed correctly by the application of the SA algorithm, described above, starting from the same set of initial parameters. The initial value equaled 1.5 for the g parameters and 300 mT for the A parameters. To the tipping angle δ for the Tl^+ perturbed $Tl^0(1)$ defect the initial value of 20° was attributed. The g parameters were allowed to vary within the interval of 0.5 to 2.0, the A parameters from 0 to 600 mT and the δ parameter between 0 and 45°. These intervals are more than large enough to accommodate the acceptable physical values of the parameters. According to the condition of a positive matrix M the maximal step size for the g parameters was 0.75, for the A parameters 300 mT, and for the δ parameter 22.5°. Our SA results given in Table 1 were obtained after more than 12000 Monte Carlo steps during which β was increased. The actual number of steps and the initial and final value of β depend on the defect to be analyzed. These results demonstrate that it is no longer necessary to

TABLE 1

The Spin Hamiltonian Parameters Obtained by the SA Procedure Are Compared to Those Obtained by an Iterative Method for Four Defects

Defect	Method	g_x	g_y	g_z	A_x	A_y	A_z	δ
Sn^+ (1) in KCl	Iterative[a]	1.819	1.788	1.959	52.9	53.8	73.0	0.
	SA	1.818	1.787	1.949	53.6	53.5	74.6	0.
Tl^0 (ortho, I) in NaCl	Iterative[b]	1.263	1.200	1.799	494.0	491.8	221.3	0.
	SA	1.265	1.208	1.778	495.6	496.9	217.6	0.
Tl_2^+ $\langle 110 \rangle$ in KCl	Iterative[c]	1.098	1.309	1.762	257.5	259.3	190.4	0.
	SA	1.097	1.317	1.763	241.3	258.1	186.0	0.
Tl^+Tl^0 (1) in KCl	Iterative[d]	1.539	1.626	1.711	12.6	14.8	240.3	32.9
	SA	1.539	1.635	1.704	17.3	13.1	240.9	32.4

Note. The hyperfine parameters are given in mT and the tipping angle δ is expressed in angular degrees.
[a] Ref. (*11*).
[b] Ref. (*12*).
[c] Ref. (*13*).
[d] Ref. (*14*).

use estimates for the initial parameters obtained by, e.g., second-order perturbation theory. This advantage of the SA method with respect to iterative methods originates from the fact that it searches the global minimum without getting stuck in a local minimum.

From Table 1 we estimate for the values obtained by the SA procedure a precision of about 1% for the g parameters, 5% for the A parameters, and 2% for the δ parameter. A higher precision for the final result can be obtained in two different ways; first, by taking the average over the accepted sets of parameters during a second run of Monte Carlo steps in which β is taken large and constant; or second, by the application of an iterative procedure, such as the least-squares-fit method in which the SA results are used as starting parameters. Both improvements will only be successful when the search of the SA procedure ended in the neighborhood of the global minimum. In that case the iterative procedure is more efficient since this algorithm guarantees a faster convergence to the minimum of function [2]. Consequently compared to the SA algorithm fewer numerical diagonalizations of the Hamiltonian [1], which are the most time demanding part of the computer program, are involved.

In our calculations such an iterative algorithm is started as soon as the function value reached by the Monte Carlo calculations is below a specified, sufficiently low value, which depends on the number of parameters to be fitted and on the number of experimental data. Using this method the parameters, given in Table 1 and obtained earlier by an iterative minimization of a function which, in contrast to function [2], needed the energy levels of the transition as input, are reproduced within the experimental accuracy given in each of the cited references.

The usefulness of the SA mainly originates from the possibility to find the global minimum of function [2] without using a second-order perturbative estimate and without the knowledge of the energy levels involved in the resonance transitions. This time-saving aspect for the experimentalist largely compensates for the larger execution

time of the computer program, especially when defects are analyzed which are characterized by large hyperfine parameters, such as the Tl^0(ortho, I) defect from our examples. Indeed, when because of the strong hyperfine interaction the zero field splitting becomes larger than the microwave energy, no trivial assignment can be given for the energy levels involved in the resonance transition and consequently second-order perturbation theory provides a bad estimate for the initial parameters. The iterative procedure then very probably yields a local minimum as solution and the calculation has to be repeated with different input data in order to find the global minimum.

From the generality of the SA procedure, it is clear that it may be used to analyze ESR spectra of different kinds because this only requires the diagonalization of the corresponding spin Hamiltonian.

ACKNOWLEDGMENTS

One of us (I.H.) is grateful to the Instituut voor Wetenschappelijk Onderzoek in Nijverheid en Landbouw for a scholarship. Another author (H.D.R.) thanks the National Science Foundation of Belgium for financial support. This work is also supported by the IIKW (Interuniversitair Instituut voor Kernwetenschappen) and the PREST Program (Ministerie van Wetenschapsbeleid) to which the authors are greatly indebted.

REFERENCES

1. G. E. PAKE AND T. L. ESTLE, "The Physical Principles of EPR," Benjamin, Massachusetts, 1973.
2. G. BREIT AND I. I. RABI, *Phys. Rev.* **38,** 2082 (1931).
3. M. R. SMITH, H. A. BUCKMASTER, AND D. J. I. FRY, *J. Magn. Reson.* **23,** 103 (1976).
4. S. K. MISRA AND G. R. SHARP, *J. Magn. Reson.* **23,** 191 (1976).
5. S. K. MISRA, *J. Magn. Reson.* **23,** 403 (1976).
6. N. TANTRAGUL AND D. BREEN, *J. Magn. Reson.* **42,** 38 (1981).
7. S. KIRKPATRICK, C. D. GELATT, JR., AND M. P. VECCHI, *Science* **220,** 671 (1983).
8. R. CAR AND M. PARRINELLO, *Phys. Rev. Lett.* **55,** 2471 (1985).
9. N. METROPOLIS, A. W. ROSENBLUTH, M. N. ROSENBLUTH, A. H. TELLER, AND E. TELLER, *J. Chem. Phys.* **21,** 1087 (1953).
10. K. BINDER, "Monte Carlo Methods in Statistical Physics" (K. Binder, Ed.), Springer, Berlin, 1979.
11. D. SCHOEMAKER, I. HEYNDERICKX, AND E. GOOVAERTS, *Phys. Rev. B* **31,** 5687 (1985).
12. I. HEYNDERICKX, E. GOOVAERTS, S. V. NISTOR, AND D. SCHOEMAKER, *Phys. Rev. B* **33,** 1559 (1986).
13. B.-R. YANG, E. GOOVAERTS, AND D. SCHOEMAKER, *Phys. Rev. B* **27,** 1507 (1983).
14. F. J. AHLERS, F. LOHSE, AND J.-M. SPAETH, *J. Phys. C* **18,** 3881 (1985).